ProLiant Essentials Recovery Server Option

technology brief, 4th edition

Abstract

The ProLiant Essentials Recovery Server Option increases system availability by allowing one ProLiant server attached to a Modular Smart Array 500 system to act as a standby for another identically configured ProLiant server. If the primary server fails, the HP StorageWorks Modular Smart Array 500 system automatically switches from the primary to the recovery server. The standby server is online in a matter of minutes—without administrator intervention. This configuration helps minimize downtime for customers with file and application servers at remote locations or branch offices where there is no onsite technical expertise.

Recovery Server Option offers customers:

- Increased system availability as a result of automated, fast server recovery.
- Fully automated server failover, ideal for unattended, or "lights out," operation.
- Ability to schedule service on a failed server at a more convenient time.
- Affordability, even for implementation at numerous operating locations.

Recovery Server Option provides fault resilience through automatic server failover. It is part of a new Adaptive Infrastructure that will help companies adapt quickly to changing business conditions, conserve valuable IT resources, and provide the highest level of customer service.

Introduction

HP is developing the tools, building the platforms, and initiating partnerships with other industry leaders to support HP's vision for the Adaptive Enterprise that will enable IT organizations to adapt quickly to changing business conditions, conserve valuable IT resources, and provide the highest level of customer service. HP infrastructure for the Adaptive Enterprise is powered by three interrelated, industry-defining technologies that are woven throughout server and storage products: continuous, secure operations; automated, intelligent management; and dynamic resource optimization.

This paper focuses on the failover technology of the ProLiant Essentials Recovery Server Option (RSO). Failover technology is a method for providing continuous, secure operations. This technology reduces exposure to unplanned events and system downtime through automated, intelligent software and hardware that allow the subsystems to predict, diagnose, and respond to potential fault conditions. Failover capability ensures that an application will be available despite a server failure.

RSO is a configuration of the Modular Smart Array 500 system to provide server redundancy capability. The Modular Smart Array 500 solution is the next-generation Ultra3 SCSI storage system designed specifically for 2-node clustering or direct-attached storage and for deployment with a range of ProLiant servers. The 4U Modular Smart Array 500 shelf houses up to 14 Universal hot-pluggable, Ultra 320 or Ultra3 hard drives, for a maximum of 2 terabyte of storage capacity using 146 gigabyte Ultra320 SCSI drives. To provide exceptional high availability, each Modular Smart Array 500 system includes one Modular Smart Array 500 controller with 128 megabytes of battery-backed cache (a second controller is supported for redundancy), a 2-port Ultra3 SCSI I/O module, and redundant, hot-pluggable power supplies with fans. The Modular Smart Array 500 system is based on standard SCSI protocols and uses a standard set of superior management tools that simplify installation and assure consistent availability.

This technology brief will explain RSO technology in detail. For information about the Modular Smart Array 500 solution, read the white paper *Modular Smart Array 500 Technology*.

Recovery Server Option overview

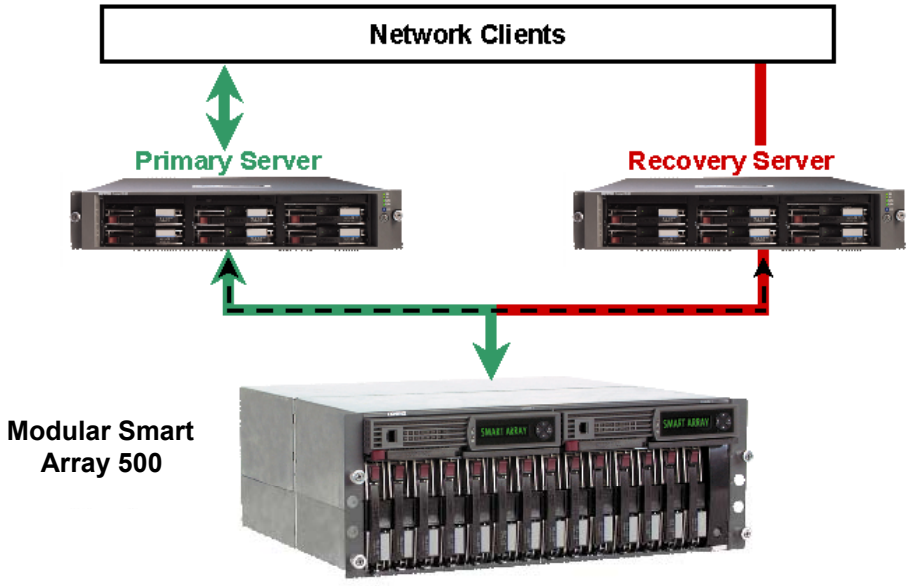
The RSO configuration comprises two identical (and identically configured) ProLiant servers connected to a common Modular Smart Array 500 system and a common network. One of the servers (the primary server) supports the network clients by default. The other server remains idle during normal use and acts as a backup (recovery) server. All hard drive storage is located in the shared Modular Smart Array 500 system. The storage system also contains a single copy of the operating system, and all the applications, drivers, and data.

When the primary server suffers a fatal hardware or operating system failure, RSO automatically takes control of the hard drives. Applications can be configured so that they automatically restart on the recovery server after RSO initializes. While the recovery server is supporting network clients, the primary server can be repaired.

To restore RSO protection to the system, user intervention is required. Once the primary server has been repaired, it must be manually reinstated as the primary server providing support for network clients.

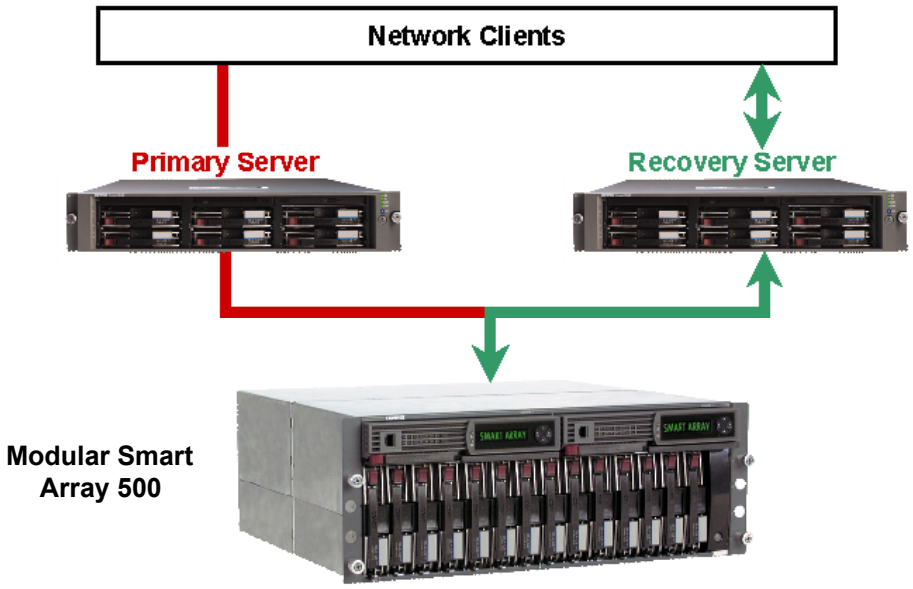
Figure 1 illustrates the RSO configuration operating under normal conditions. The primary server supports network clients. The recovery server monitors a heartbeat signal (dashed line) that the RSO software periodically sends from the primary server through the storage system. As long as the recovery server continues to receive heartbeats according to schedule, it remains idle.

Figure 1. Normal RSO operation.



If the recovery server does not receive a heartbeat within a preset time period, a failover begins. The recovery server commands the storage system to disconnect from the primary server and to link to the recovery server (Figure 2), which then boots the operating system. If Insight Manager is running on the system, it sends an alert to the system administrator that a failover has occurred.

Figure 2. RSO configuration after a failover



Network clients experience a temporary service outage until the operating system on the recovery server becomes functional. Since both servers boot from the same operating system image, they have the same logical network identity. Hence, network clients can log on and regain access to the data files as normal when the failover is complete.

When the cause of the failover has been corrected, the administrator powers down the system and manually restores the RSO configuration. In most cases, restoring the configuration involves either reinstating the primary server after it has been repaired or replacing the primary server.

The automated RSO process eliminates the need for intervention by the system administrator to recover from a server failure; therefore, it increases system availability and offers an excellent solution for unattended operation. It offers minimum downtime for customers with file or application servers at remote locations or branch offices where technical expertise may not be available on site. Moreover, the system administrator can schedule maintenance of the failed server at a convenient time, such as at the end of the business day or during a period of relatively light activity. RSO is an effective, affordable solution for highly available access to applications and data—even for companies with numerous locations.

System requirements

RSO supports several combinations of ProLiant servers connected to Modular Smart Array 500 systems. Server hardware configurations should be identical to prevent operating system configuration problems when a failover occurs. The primary and recovery servers boot from the same Modular Smart Array 500 system. The servers should have identical processor configuration, and the network interface controllers (NICs) and SCSI controllers must be installed in the same slots in both servers.

To ensure compatibility, RSO configurations must fulfill the following requirements:

- The same model of server must be used for both the primary and recovery servers, and they must be configured identically.
- Only one array controller in each server can be connected to the storage system. Each controller in one server must be matched with an identical controller in the other server, installed in the same slot number, and configured identically.
- The NICs on the two servers must be of the same type, installed in the same slot number, and configured identically.
- Hard drives must be installed only in the storage system. Hard drives cannot be connected directly to either server.

For detailed information about the supported storage system, server models, and array controllers, refer to www.hp.com/products/sharedstorage.

Operating system

RSO operates independently of operating system and application software, but it requires the appropriate software drivers. RSO supports these operating systems:

- Microsoft® Windows® 2000
- Microsoft Windows NT® 4.0
- Red Hat Linux® 7.2, 7.3, 8.0
- Red Hat Enterprise Linux® AS 2.1, ES 2.1
- SuSE Linux Enterprise Server 7
- Novell NetWare 6.0
- Novell NetWare 5.1

To identify the specific operating system versions that RSO supports, check the RSO QuickSpecs at www.hp.com/products/sharedstorage.

Insight Manager support

Insight Manager is a server management tool that provides intelligent monitoring and alerting, remote maintenance, and visual control of servers in a network environment. Insight Manager supports the RSO configuration by alerting administrators if either server fails.

Use of Insight Manager with RSO is highly recommended. For customers using Insight Manager, two alarms are associated with RSO. These alarms appear as messages on the Insight Manager console screen.

The first alarm is sent by the recovery server. It indicates that the primary server has failed, a failover to the recovery server has occurred, and the recovery server has booted the operating system.

The second alarm is sent by the primary server. It indicates that the primary server is no longer receiving an acknowledgment message from the recovery server. Loss of the acknowledgment message indicates one of the following conditions:

- The recovery server has failed.
- The recovery server has been powered off.

If any of these alarm messages appears on the console screen, the system administrator should schedule service to repair the failure condition and restore the recovery server to its original configuration.

For more information about Insight Manager, go to www.hp.com/servers/manage.

How RSO works

In the past, RSO was implemented using both software and hardware. The RSO software worked with Insight Manager to monitor the configuration and alert the administrator to failures. In addition, RSO required use of a Recovery Server Interconnect and a Recovery Server Switch. The Recovery Server Interconnect was a serial cable that connected the primary and recovery servers and transmitted a heartbeat signal from each server to the other server. The Recovery Server Switch was an intelligent SCSI switch installed in each switchable ProLiant Storage System that automatically switched the SCSI bus between the two servers.

Now, RSO is implemented entirely in software. RSO is enabled and configured entirely through the Option ROM Configuration for Arrays (ORCA) utility. Once RSO is enabled, the configuration process is completely automated, with intuitive prompts to lead the administrator through the process.

The primary server must be connected to the SCSI port on the Modular Smart Array 500 system, and then RSO must be enabled and configured. Once RSO is enabled and configured on the primary server, the recovery server must be connected to the second SCSI port on the Modular Smart Array 500 system. Then RSO must be configured on the recovery server.

The heartbeat signal is transmitted through the SCSI links to each server. As part of the RSO configuration, the administrator will set a timeout value for the heartbeat signal. ORCA on the recovery server monitors the heartbeat signal sent by the RSO software on the primary server. In the event of a failure on the primary server, the Automatic Server Recovery (ASR) will sense the lack of operating system activity and reset the primary server. When the heartbeat signal fails to arrive in the set timeout interval, ORCA on the recovery server will issue a command that disables reads and writes from the primary server, essentially blocking the primary server from accessing the Modular Smart Array 500 system. ORCA then exits and allows the recovery server to boot normally.

The primary server will remain dormant, awaiting service. It can then be disconnected and removed for hardware repairs while the recovery server is running the operating system. The recovery server must be brought down in order to reinstall the primary server. The administrator can schedule this reinstallation so that it has the least impact to users.

Application behavior and data loss

The chance of data loss is lowest with applications that behave predictably in the event of system failure.

Database

Most database applications use the concept of committed transactions to ensure data integrity if a system fails. Databases such as Microsoft SQL Server and Oracle® Workgroup Server perform automatic recovery operations when they are restarted after a server failure. This recovery operation ensures the integrity of the database and removes transactions that were not committed or completed when the system failed.

File and print

Applications that use a ProLiant server as a file or print server will operate the same way that they do on a single server. The risk of data loss caused by server failure remains the same. With RSO, however, clients can regain access to applications and stored data more quickly.

Planned shutdowns and forced failovers

RSO provides an option that allows administrators to plan shutdowns and force failovers. When the F8 key is pressed on the recovery server, a manual failover is initiated. Once the failover to the recovery server is complete, the administrator can shut down the primary server and perform any necessary maintenance. Or, in the case of a forced failover, it can be initiated without shutting down

any components of the system. A forced failover to the recovery server can be initiated at any time from the console of the recovery server.

Power loss to the storage box

If power is lost to the Modular Smart Array 500 system in an RSO configuration, the primary server will not boot in an unattended manner when the power is restored. A failure of this type will cause network clients to lose service. When the primary server is powered on, the administrator will be prompted to run diagnostics or to continue a normal boot sequence. This illustrates the importance of an uninterruptible power supply.

If the system is unattended when the power is restored, and the servers boot before the storage system is fully initialized, neither server will see the logical volumes on the storage system; and clients will remain without service until the administrator initiates a normal boot sequence.

Failover speed

The total time required for the recovery server to assume the function of the primary server is the sum of the following factors:

- The time that elapses between the moment of actual failure and the moment at which the recovery server fails to receive a scheduled heartbeat
- The time required for the operating system to boot, which depends upon:
  - The capacity and number of hard drives that are in the storage system
  - The time required for applications to open and perform integrity checks on their files (which in turn depends upon the application and on the size, type, and number of files).

Failure detection

Any event in the primary server that stops RSO from generating the heartbeat causes a failover (see Table 1). Examples of such events include:

- Catastrophic and unrecoverable hardware failure in the primary server, such as loss of the processor or uncorrectable memory errors
- Failure of the primary server power supply
- Failure of the operating system

Table 1. Possible failure scenarios

Scenario	Results
Primary server fails	Recovery server takes over network support
Recovery server fails while the primary server is supporting network clients	Insight Manager sends an alert, reporting that the recovery server has malfunctioned
Recovery server fails while the primary server is offline	Network clients lose service
Storage system fails	Network clients lose service

The following events on the recovery server also cause a failover:

- The recovery server does not receive heartbeats, although the primary server is sending them.
- The F8 key is pressed on the recovery server to manually begin a failover.

A failover does not occur if the recovery server still receives heartbeats, despite the component failure. For example, if the NIC on the primary server fails, the primary server cannot connect to network clients. However, RSO can still send the heartbeat to the recovery server. Insight Manager detects most failures of this type, and it should be left running for this purpose.

Generally, any failure that is detected by ASR is detected and acted upon by the recovery server. However, some failures will cause the primary server to malfunction without causing loss of the heartbeat message. For example, failure of the NIC could render the primary server unusable, but the RSO driver would still send the heartbeat message to the recovery server. The recovery server cannot detect failures of this type; therefore, an automatic failover will not occur.

Servicing the failed server

To reestablish RSO protection after a failover, the failed primary server must be repaired or replaced and brought back online. RSO makes it possible for the system administrator to schedule service on the primary server (on or off site) at a more convenient time while the recovery server is active.

Once the failover occurs, no drives are electrically attached to the disk controllers in the primary server. This may limit diagnostic activities that can be performed on the failed primary server. By disconnecting the primary server and adding other drives to the primary server, full diagnosis can be performed on the failed primary server while the recovery server is running.

Restoring configuration after failover

After the primary server is serviced, the original configuration can be restored. The recovery server must be power cycled to reinitialize the RSO configuration. The disk drives will be electrically connected to the primary server, which will boot the operating system. The recovery server will return to its role of listening for the heartbeat message from the primary server.

Client behavior

When a failure of the primary server occurs, clients attached to the network experience service interruption. The symptoms experienced by the clients vary depending on the operating system, the network protocol, and the application.

Windows

For Windows clients, failure of the primary server appears as the inability to read from or write to the network device. In most cases, after the recovery server has booted, client connections to the recovery server will resume automatically.

The primary and recovery servers both boot off the same set of storage disks. For this reason, after the recovery server boots and becomes operational, the Windows Event Log may show entries that were made when the primary server was active. Likewise, after the primary server has been restored to active service, the Windows Event Log may show entries made while the recovery server was active.

Linux

For Linux clients, failure of the primary server appears as the inability to read or write to the network device. In most cases, after the recovery server has booted, client connections to the recovery server will resume automatically.

NetWare

For NetWare clients, failure of the primary server appears as the inability to read or write to the network device. During a failover, the behavior of a NetWare client is not predictable. In some cases, the connection to the server could be maintained through the failover; in other cases, the connection could be lost. If the connection is lost, server functionality will be lost.

If a failover occurs, application software should be exited in the normal fashion and NetWare clients should be rebooted.

Disk integrity

Failure of the primary server can be caused by several different conditions ranging from software faults in the operating system to hardware failure. Depending on the nature of the fault and the disk activities occurring at the time of the fault, the disk data structures may be corrupted and may require corrective processing before the recovery server boots the operating system.

Windows

The NT file system, NTFS, is required for all Windows disk partitions. Additionally, it is recommended that the Windows system disk and other executables be placed on a separate Windows disk partition. Other partitions should be used to contain data.

Linux

For Linux, disk checking can be configured in the File Structure Table, FSTAB. For the boot volumes, this is always done at system boot. It is recommended that customers use a robust journaling file system.

Netware

NetWare automatically performs an integrity check of volumes that it mounts. However, it is recommended that the SYS volume be configured primarily for use as storage for executables, not for data files.

Conclusion

The ProLiant Essentials Recovery Server Option is an effective high-availability solution for providing fault resilience for business-critical applications. The automated RSO process eliminates the requirement for intervention by the system administrator to recover from a server failure. RSO increases system availability and offers an excellent solution for unattended operation. RSO ensures minimum downtime for customers with servers at remote locations or branch offices where technical expertise may not be available on site.

Call to action

Please direct comments regarding this communication to the ISS Technology Communications Group at this Internet address: TechCom@HP.com

TC031104TB, 11/2003